NETWORK MANAGEMENT ARCHITECTURE

Background of the Invention

The present invention is directed to the field of network management. The invention has particular applicability to the improvement of a desired performance metric in a wireless communication network. A typical wireless network includes a plurality of wireless clients such as laptop computers and cellular phones that exchange data packets with a network backbone via a radio card. These clients gain access to the network backbone by way of an access point (AP), such as a wireless base station, connected to the network by a hardwired connection or a fixed Ethernet connection.

The advantage of a wireless network is that clients are mobile units that can vary in position (such as with wireless laptop computers) or can even be in motion while in use (such as with cellular telephone service.) However, these wireless clients have only a finite range wherein the base station may be accessed. Beyond that range, a client cannot maintain a connection to the base station. Consequently, in order to cover a large area, a number of such access points are required so as to maintain service to the clients. At any given time, an AP serves a number of clients in a unit called a Basic Service Set (BSS) and the region served is called a Basic Service Area (BSA).

As mobile clients vary in position, each BSA has an ever changing topology of wireless connections. However, due to regulatory frequency allocations, there is not enough available spectrum to insure that each client can gain access to the network at the same time. A number of schemes are known to permit clients to share access through an AP within a BSA. However, it may happen that multiple BSA's can be on the same channel, producing interference from "packet collisions" that break the network, resulting in "packet drops". Other reductions in

5   quality-of-service (QoS) can occur such as packet delay and packet jitter, which can also lead to

packet drops.

In order to reduce interference from packet collisions in a wireless network, access can be

controlled using a distributed coordination function (DCF), which relies on "carrier sense,

multiple access collision avoidance" (CSMA/CA). With DCF control, each AP and client radio

10   listens to the air and waits a random amount of time for the air to clear. If the air is then clear,

the radio transmits a packet. In this manner, network access is regulated through statistics. This

technique is found to statistically provide a modest reduction in packet collisions, though not

altogether preventing collisions. The principle drawback is that only one BSA can be on the air

at a time, since frequency bandwidth is restricted by FCC regulations. Consequently, the

15   network throughput of BSA's sharing the air is fixed and is divided by the number of radios on

the common channel. As the number of AP's and clients sharing the air is increased, access is

further diluted, resulting in increased packet collisions and a significant waste of bandwidth.

As shown in Fig. 1A a typical DCF control uses an AP with an omni-directional antenna

10, which transmits and receives in a 360 degree radiation pattern to access all the associated

20   clients. As shown in Fig. 1B, the problems with DCF control are partially ameliorated by the use

of an adaptive directional antenna 20 having a narrow radiation pattern 22 that can be varied so

as to be selectively directed to each client in the BSA. The adaptive directional antenna 20 can

be an array of antennas with rapid switching therebetween.

This adaptive directional antenna 20 can be used in a point coordination function control

25   (PCF) to perform time division, multiple access (TDMA) rather than CSMA/CA. In fact,

adaptive antennas are better used by PCF, which offers improvement in QoS. With PCF, every

client 24 is assigned a time slot by the AP in a predetermined manner in which the AP queries

2

5   the clients, and receives a packet or transmits a packet. PCF is superior DCF in that access can

be regulated and prioritized to reduce packet drop outs from packet collisions and other QoS

problems resulting from queuing issues, such as packet delays and packet jitter. PCF control

works especially well for real-time communications such as voice traffic.

However, PCF control creates complications in a wireless network where multiple BSA's

10   share the same channel. AP's work independently, with no communication between each other.

Consequently, interference can still occur between radios in different BSA's, especially if they

lie along the same line of sight. Fig. 1C shows adjoining BSA's A, B having respective AP's

20a, 20b, each connected to a network backbone 30. The AP 20a of BSA A can be in contact

with a client 24a while in line of sight with a client 24b associated with the BSA B. Interference

15   can result if both cells (AP's and clients) become active at the same time, especially if e.g. AP

20a transmits to client 24a while client 24b transmits to AP 20b. As the clients are mobile, the

network topology is always changing.

Summary of the Invention

In view of the difficulties and drawbacks associated with previous network architectures,

20   there is therefore a need for a network management system that reduces interference between

BSA's in a PCF-controlled network.

There is also a need for a network management system that improves quality of service

factors.

25   There is also a need for a network management system that increases network throughput.

3

5   These needs and others are satisfied by the present invention in which a network includes a network backbone and one or more macro service areas. Each macro service area includes one or more basic service areas, which each include an access point and a plurality of associated clients. A management system is employed for connecting the access points to the network backbone. The management system regulates access of the clients to the network so as to control

10   a desired network parameter, which can relate to throughput, quality of service, or any other desired network parameter.

As will be realized, the invention is capable of other and different embodiments and its several details are capable of modifications in various respects, all without departing from the

15   invention. Accordingly, the drawing and description are to be regarded as illustrative and not restrictive.

Brief Description of the Drawings

Figs. 1A, 1B and 1C respectively show AP's with an omni directional antenna, an adaptive directional antenna, and a pair of adjoining BSA's in a network.

20   Figs. 2A and 2B depict a macro service area in accordance with the present invention.

Fig. 3 depicts a first embodiment of a management system according to the present invention.

Fig. 4 illustrates an extended multiple MSA network using the management system shown in Fig. 3.

25   Fig. 5 illustrates a multiple MSA network using a second embodiment of the management system according to the present invention.

4

Figs. 6A and 6B are flow charts indicating alternative methods for implementing the present algorithm.

Fig. 7 depicts multipath and angle-at-arrival as employed with the present method.

Detailed Description of the Invention

As shown in detail in the figures, a wireless network is disclosed having a centralized control for managing traffic of client data packets through the network. In this way, optimized network operation is maintained, to improve throughput and/or another performance metric such as the factors related to improving quality of service.

As shown in Fig. 2A, a wireless network 100 includes adjoining BSA_1 and BSA_2, potentially having at least a portion of channel overlap (i.e. frequency and spatial). Each $BSA_{1, 2}$ includes an access point 110a, 110b that communicates with clients 114a, 114b, preferably using an adaptive directional antenna or antenna array in a PCF control scheme. A management system 120 is employed to connect the AP's 110a, 110b to the network backbone 130, in order to define a macro service area (MSA). In a MSA, the management system 120 controls the BSA's so as to minimize co-channel interference. In an MSA, the management system 120 instructs the AP's 110a, 110b to regulate client access in an efficient manner for specific network topologies, loads and QoS requirements. An MSA can manage many network variables, including time division, buffering, bandwidth, frequency and space. In the preferred embodiments, an adaptive control algorithm would be employed that would manage these network variables in view of the changing parameters and requirements of the network. It should be noted that Fig. 2A shows an MSA serving only two BSA's, with a small number of clients. It should be appreciated that this

5

5   and the other figures are only given for illustrative purposes, and any number of BSA's and/or

respective clients could be contemplated without departing from the invention.

In order to coordinate and prioritize network traffic from a number of clients in different

BSA's, the algorithm of the management system 120 operates with a goal function to optimize a

desired network metric. As shown in Fig. 2B, each BSA can be assigned a unique coordinate

10   system. In principle, when a client associates with a particular BSA, the respective AP 110a,

110b calculates the directions of arrival for the client signal with respect to its coordinate system.

The management system 120 must transform those coordinate systems to its own MSA

coordinate system, e.g. to reference the locations of origins and rotation of the reference axes.

However, due to limiting factors such as measurement error, perfect coordinate transformation

15   data may not be obtainable. Thus, the algorithm must be adaptive in order to operate with

imperfect or transient information.

Preferably, the algorithm functions to control a particular desired network parameter.

Specifically, the algorithm preferably operates with a goal function or metric that seeks to

approach an extremum of a network parameter. For example, in a DCF network, using an

20   omni-directional antenna, it would be sufficient to use temporal access techniques to manage

time. The algorithm may optionally direct the AP's 110a, 110b to control access of clients to the

channel so as to maximize network throughput. In such a case, the goal function is the network

throughput in bits per second and the algorithm seeks the assignment of clients to appropriate

time slots in such a manner that the throughput is maximized. In operation the algorithm would

25   determine the clients that can be on the air at the same time by observing previous network

packet collision history and adjust the client to specific time slots. As shown in Fig. 6A, this

method could be implemented by acquiring collision rate and throughput data from the APs in a

5   MSA. If poor throughput is determined, the algorithm adjusts client access to the network to reduce the collision rate.

It may also be desirable to manage time and space in a PCF network using an adaptive antenna. A wireless network of four AP's each having a throughput of 10Mbps each AP could be simultaneously accessing clients by managing space via an adaptive antenna, with respective

10   client time division multiple accesses on each AP selected to avoid simultaneous access between potentially interfering clients.

In another case, the algorithm would use packet angle-of-arrival information to determine which particular clients could access the channel at the same time using an adaptive antenna array and beam/null forming methods. (Beam/null forming is the computation of orthogonal

15   (non-interfering) antenna array patterns which can be used by the AP to place more than one client on the channel at the same time on the same frequency.) The goal function could be the maximizing of network throughput or minimization of packet jitter. As shown in Fig. 6B, the . method can be implemented by acquiring angle-of-arrival (AOA), RSSI and throughput data from the APs in the MSA, and adjusting client access to the network in response to these factors.

20   By using an adaptive antenna to manage space and time, network throughput would approach the maximum 40Mbps, as compared with only 10Mbps available on a competing DCF network.

Current wireless network trends indicate increasing demands for "real-time" wireless service, e.g. cellular telephone voice communications. Thus, the algorithm of the present invention may operate with a metric that seeks to improve one or more quality of service

25   parameters. For example, queuing between multiple AP's may result in packet delays. This can result in packet drop out in a real-time voice message. Also, packet jitter may occur if packets arrive too soon or too late. The present management algorithm may be tailored to a metric that

5   seeks to minimize packet jitter and packet delay by granting priority access to real time packets, resulting in the timely sequential arrival of those packets, thereby improving network QoS.

The present algorithm of the network management system 120 could be realized by a number of implementations. For example, an associative neural net could be used, or a root mean square error program could be used. Also, a number of artificial intelligence schemes

10 could be employed. Some exemplary implementations are given as follows. In one example, we assume a simple case where AP's are on a regular grid, AP's are oriented on a regular grid, preferably employing TDMA operation. Using an antenna system whereby it is possible to service a particular area around a given AP, consider the positional topological arrangements for covering the total service area given the requirements that it is desired that more than one client

15 be serviced at a time from each AP at any time slot. This means that by monitoring the positions for a given placement of AP's and clients for a desired type of antenna technology, there will be a finite number of topological arrangements that can be constructed that permit, on average, the AP's in the network to allow network access more than one client per TDMA time slot. These distinct topological arrangements could be preprogrammed into the main management program if

20 the AP's are oriented and placed on a regular grid. Access to the network can then be regulated for each topology to improve network performance, e.g. maximize network throughput.

In a real system where the RF channel is less than ideal and there are significant amounts of multipath, some topological arrangements would be unusuable. This cannot be determined from the geometrical information provided by the placement grid and the AP orientations.

25 However, this could be derived from direct measurement or by trial and error as the network is in operation. In the latter case the main controller could conduct a network calibration routine. During this phase the main controller could instruct a particular AP to transmit in a particular

5   direction while all other AP's "listen" (or detect the signal). The listening AP's would record the

received signal strength and direction of arrival of the signal and report them to the main

controller. This process would be repeated until all AP's and the paths between them were

calibrated. This could also be done for the associated clients within the network. They also

could be constructed to both transmit and listen and report their information back to the main

10   controller. After this process the main controller could fairly accurately determine the allowable

topological configurations. Well known algorithms such as "MUSIC" or "ESPRIT" could be

used to calculate the "angles of arrival" (AOA's). As shown in Fig. 7, each client associated

with an AP has well defined signal propagation paths between itself and the AP, each having a

specific angle of arrival with respect to the AP or client, which ever is in receive mode. One

15   direct and two reflected paths are shown for transmission from client to AP. Each path

approaches the AP along a unique direction. These angles are the angles of arrival.

The above mentioned calibration process could also be used on a system where the AP's

were neither oriented nor placed in positions known to the main controller. The calibration

procedure could indicate relative directions which along with the relative received signal strength

20   and path delays could provide a reasonably accurate picture of the actual placement of the AP's.

This could be accomplished by Joint Angle and Delay Estimation (JADE).

Over a period of time, the configuration of the network would change and some

topological arrangements would no longer be viable. In this case the main controller would have

to modify the management routine and select a modified set of topologies to assign the various

25   TDMA time slots. The information coming into the main controller could consist of several

items, in particular, orientation of clients being serviced relative to the AP to which it is

5 associated, priority of traffic flowing to and from the clients, the level of performance provided

to the client, including:

   a. Data rate used;

   b. Numbers of retries;

   c. The topological configurations used when talking to the client; and

10   d. Other items such as would occur to those skilled in the art.

By knowing the topological configurations used while servicing a particular client the

main controller could try other configurations that would possibly eliminate the new multipath

interference caused by having certain clients on the air at the same time.

The present MSA management system 120 can be implemented in several ways, two

15 implementations are contemplated hereinbelow. In a first preferred embodiment, the

management system 120 is an enhanced Ethernet switch. In the second preferred embodiment, a

particular AP is designated as the MSA controller 120 for a specifically assigned group of

BSA's. A general switch architecture of the first preferred embodiment is shown in Fig. 3. Fig.

4 shows the overall network architecture using this switch. A high speed Ethernet media access

20 controller (EMAC) 140 is connected to a wired network backbone 130. Several additional

EMAC's 142 are provided for sending and receiving data from AP's in the managed BSA's.

Additionally, there are provided dual multi-level queues (DMLQ) 144 for each EMAC data path

to handle incoming and outgoing packets of varying priority according to the system QoS

requirements. Lastly, a controller 146, which can be a Power PC or another type of dedicated

25 processor, would regulate the data flows between the network backbone 130 and the controlled

AP's. To fully support the more advanced QoS methods the switch 120 would preferably handle

5   Reservation Protocol (RSVP) and its extension, subnet bandwidth management (SBM) as well as

be 802.1p and 802.1Q compliant.

For such an arrangement it is possible to accomplish complete control over the AP's

connected to the switch 120. For purposes of clarity, it is useful to assume that all the BSA's in

this MSA are on the same frequency. As shown in Fig. 4, such a situation might arise if several

10   MSA's (MSA_1, MSA_2) were required to cover a particular enterprise installation and all available

channels were assigned to the different MSA's to prevent them from interfering with one

another. In such a configuration, all BSA's in a respective MSA would interfere with one

another if permitted access to the channel at the same time. This may not be the case in all

environments since it may be possible for more than one AP to access the channel at one time by

15   proper channel allocation. Under these conditions each respective switch 120_1, 120_2 would

control the timing of when each AP (110a, 110b, 110c, 110d, 110e, 110f) accesses the channel.

In the case where all BSA's have the same QoS requirements and similar traffic patterns all that

would be necessary is a sequential servicing of each respective BSA. In the case where QoS

needs were unequally divided between BSA's the switch could effectively adapt since it has

20   direct control of the priority queues for each AP and is completely informed as to the QoS

requirements of each BSA from the state of the priority queues. Due to centralized control and

complete knowledge of network resource usage, the switch 120 could accurately allocate

resources and determine when to admit new data streams. In the case of a very large enterprise

or very large public WLAN installations it may be necessary to implement many MSA's, some

25   of which may be on the same frequency and interfere if permitted to access the shared medium at

the same time. In this case another level of management would be needed. A similar switch that

11

5   runs a modified control algorithm could provide this functionality. This could actually be

accomplished with the same product jumpered or programmed to manage MSA's.

Referring now to Fig. 5, a schematic of the network architecture is indicated for

designated AP control of an MSA. In this implementation, a unique AP within an MSA is

selected to be the MSA controller. During operation, this control AP 120_1, 120_2 (CAP) manages

10   its own traffic in addition to the traffic of all other AP's within its respective MSA. All priority

buffering is done at each AP within the MSA, hence the CAP 120_1, 120_2 is not directly aware of

the QoS needs of its MSA. The CAP is informed of the network load and QoS requirements and

manages resources within the MSA by receiving requests for extra bandwidth from the various

AP's which it controls over the existing network backbone 130. The CAP 120_1, 120_2, based on

15   the current MSA status, would either allow or deny access to the channel. This is done by

changing the times and rates at which the controlled AP's could access the channel. Note that all

AP's within the MSA must be informed of the new channel access times and service rates, as the

variables change in a transient, mobile network.

By controlling the network as discussed above, many possibilities arise that can

20   significantly increase network performance. In the above, we have discussed the ability of the

MSA controller to control at what time and for how long and at what rate the AP's can be

serviced. This level of control significantly improves the QoS when coupled with RSVP/SBM

and 802.1p/Q. However, it is also possible to control frequency and space if more complicated

network architectures are admitted, particularly if adaptive directional antenna control is

25   provided at the AP's. This would be especially useful when all BSA's within a given MSA must

be on the same frequency. In this case, "spatial domain multiple access" (SDMA) would be

12

5   possible and would permit more than one AP access to the channel at a time. This could have a

dramatic impact on network throughput and hence QoS.

As described hereinabove, the present invention solves many problems associated with

previous type systems. However, it will be appreciated that various changes in the details,

materials and arrangements of parts which have been herein described and illustrated in order to

10   explain the nature of the invention may be made by those skilled in the area within the principle

and scope of the invention will be expressed in the appended claims.